Chapter 1— Overview

1.1 Preface

This document provides you, the DoD program manager or M&S (Modeling and Simulation) manager with an understanding of basic Verification, Validation, and Accreditation (VV&A) terminology and techniques. Its goal is to help you develop an informed and independent judgment about how credibly models and simulations (M&S) are being integrated into your program. To understand why you should be concerned with the material in this document, imagine yourself in the following situation.

You are a senior officer or civil servant working for one of the Services. You have just been tasked to provide a comprehensive solution to a major military problem. That problem may be the development of a new weapon system, the design of a training exercise, or perhaps the definition of military force structure requirements in your branch of the Service for the next three decades. You have little time, less money, and only meager human resources to complete the task. You know (or you have heard) that one of the ways to save time, money, and human resources is to take advantage of the breathtaking array of models and simulations that have been made possible by the dramatic increase of computer hardware and software capability in the last decade. You don't know much about M&S and perhaps still less about particular simulations, but you know how to get to the people who do. So you set up an M&S shop within your organization; you allocate precious resources to a staff of analysts, scientists, engineers, and warfighters; and you charge them with the delicate task of pulling together a credible M&S effort that will meet or support key program objectives while saving time and money. You figure if these people can't do it, nobody can. You walk away happy.

Time passes. Things go along pretty well for a while, or at least they appear to. Every so often you call for a program review that includes the status of M&S efforts. "Everything's fine," you're told. The M&S suite has *been selected and stabilized, M&S outputs have been related to key program Measures of Merit (MOMs), M&S reviews have been scheduled and conducted, and your M&S shop is confident that their results are credible on the basis of "VV&A."[1] "VV&A?" you ask. "What's that?" Your M&S team throws alarmingly technical terms around that make it sound as if big money is being spent. "Not to worry," they say. The "industry" has been doing "VV&A" for years; the Military Operations Research Society (MORS) has standard definitions for key technical terms and techniques; and no unnecessary "V&V" is being done. The models and simulations supporting this program will be "VV&A'ed" in time to meet major program milestones. You walk away happy. Sort of.*

You do a little research. You discover that, far from being a compact, tightly knit, well-defined discipline, VV&A spans a broad spectrum of activities. You discover that the depth and breadth of these activities depend not only on the kind of M&S to which they are applied but also to the specific application for which the M&S will be used. You discover that "community consensus" about the definitions of verification, validation, and accreditation exists at only the most general level. You also find out that the definition of the V&V techniques that should be used for specific types of models and simulations and how these techniques should be applied to establish the credibility of M&S when used for particular applications is a subject of intense debate. You discover that a major high-level review of your program is fast approaching, and you suspect that some questions about all this VV&A business will come up because of the attention given it in recent DoD and Service policy documents. You wish you knew how to make an independent judgment of how well your M&S team has met its critical milestones to support your program's objectives. You walk away maybe not so happy.

Sound familiar? Then this document is for you.

The information in this document has been compiled from a wide variety of sources, including recent DoD Directives and Instructions related to M&S management and VV&A; software industry standards and practices; the practical experience of numerous ongoing VV&A efforts across the DoD and industry; academic texts and professional literature; and professional societies and organizations intimately familiar with M&S and VV&A. The hope is that this broad array of experience, concisely presented, will encourage you to pursue VV&A of M&S with confidence, vigor, and insight.

In addition to this introductory section, this chapter consists of six sections that provide (a) an understanding of basic V&V techniques and terminology (Section 1.2); (b) an appreciation of the value of VV&A (Section 1.3); (c) a discussion of where VV&A fits in the scheme of M&S (Section 1.4); (d) a discussion of limitations to VV&A (Section 1.5); (e) a general introduction (Section 1.6) to some practical aspects of VV&A, such as tailoring V&V tasks to the requirements of your specific application, who should be doing what (and why), and costing and scheduling considerations; and (f) a description of the rest of this Guide (Section 1.7).

1.2 What Is VV&A?

This section defines these terms: model, simulation, simulator, M&S, verification, validation, accreditation, and other related terms.

M&S credibility is measured by verification and validation (V&V) and formally approved as adequate for use in a particular application by accreditation. The entire process is known as VV&A. Before we define the individual elements of VV&A, let's get a few preliminary terms out of the way.

1.2.1 Terminology

One of the most confusing aspects of M&S terminology is the difference between a model and simulation. In fact, many people in the M&S community either do not really know (or do not really distinguish) between the two in conversation. In fact, there is no official consensus as to the definitions of these terms, nor do we propose to settle the debate within the context of this venue. The general distinction between a model and a simulation will be important, however, when we talk about the details of VV&A. We have developed an approach to explaining the terminology, therefore, that is consistent with (most) current definitions, has practical utility, and is not illogical.

According to DoDD 5000.59, a *model* is "a physical, mathematical, or otherwise logical representation of a system, entity, phenomenon, or process." A *simulation* is "a method for implementing a model over time." But what do these definitions mean in lay terms? And how does the distinction between them affect the nature of VV&A activities?

A *model* is a conceptualization, an abstraction of some physical phenomenon or process into mathematical equations and solution approaches (called "algorithms"), each with its own assumptions, limitations, and approximations. For example, the radar range equation is a model, an abstraction of the radar detection phenomenon into an equation that makes certain assumptions about how radar energy interacts with targets, clutter, and the atmosphere. If you take this equation and convert it into a computer program (software) to solve it for particular scenarios as a function of time (say, to determine the detection history of a combat aircraft during a mission from a fixed radar site), the result is a *simulation*, which is a software framework that executes a model (or models or model pieces) in the proper order, provides timing and coordination between them, and controls the inputs and outputs. Thus, a model is an abstraction of a phenomenon into mathematical equations and algorithms, whereas a simulation is the software implementation and solution of those equations and algorithms over time within the context of a scenario. A model can exist without a single piece of software; a simulation is the software that implements the model over time.

Another potential point of confusion is that many people in the M&S community use the term *M&S* to stand for both *models and simulations* and *modeling and simulation*. *Modeling and Simulation* is an analytical problem-solving approach; *Models and Simulations* are mathematical abstractions and software implementations. Although the community uses the terms interchangeably, this document distinguishes between the two.

And, lest this topic become too easy to understand, we introduce yet another commonly used term that may cause the newcomer to M&S terminology some confusion: *simulator*. In its broadest sense, a simulator is a training device made up of some combination of hardware and software designed to provide an artificial (but suitably realistic) environment in which a human player can interact with those aspects of reality in which training is desired and within which all aspects of reality that are simulated interact realistically with each other. Flight training simulators come to mind as good examples. Not all aspects of reality need to be simulated in a simulator, only those crucial to the goal of training. Similarly, not all aspects of reality that are simulated need to be simulated with complete realism, only enough to ensure that training objectives are met.[2]

Simulators also can be used for testing, but here their required level of realism typically is greater. The most obvious case of simulators built for both training and testing applications are the open-air threat missile system simulators found on many DoD test ranges. These systems are used not only to train pilots in the proper use of available equipment and combat tactics but also to test the effectiveness of new electronic countermeasures (ECM) systems designs. With these simulators, the environment created is not enclosed (as it is in the case of a flight simulator), but the simulator still reproduces those aspects of reality essential to the training or testing application (e.g., a realistic, open-air RF environment).

One final point concerns the distinction between the terms *verification* and *validation*. Most people have an intuitive understanding of the meaning of the term "validation" with respect to M&S. Ask them to tell you the difference between verification and validation, however, and you're likely to get a blank stare, because these two words have the same or similar meanings to many people. To the M&S community, however, there are important distinctions.

The distinction between the two is most easily remembered in terms of their focus. At the risk of oversimplification, verification focuses on M&S *capability*, whereas validation focuses on M&S *credibility*. *Verification* ensures that a simulation meets all the requirements specified by the user and that it implements those requirements correctly in software; *validation* ensures that a simulation conforms to a specified level of accuracy when its outputs are compared to some aspect of the real world. We'll explore the nuances associated with determining the level of accuracy required of a simulation later. For now, just realize that verification and validation add separate, distinct, and essential kinds of credibility to M&S. Neither achieves its fullest contribution to M&S credibility without the other.

With basic definitions and distinctions out of the way, let us now turn to more detailed descriptions of verification and validation (V&V).

1.2.2 Verification Basics

According to DoDD 5000.59, *verification* is "the process of determining that a model implementation accurately represents the developer's conceptual description and specifications." In more colloquial terms, verification is the process of determining that a *model* and its resultant *simulation* (remember our definitions) accurately represent both what is required and what the M&S developer says will be built for you in accordance with those requirements.

If you are planning on developing models or simulations for use in your program, you need to do two things before a single line of software (usually referred to as *code*) is written. You need to build and verify a *conceptual model* from which the code will be written (Conceptual Model Verification), and you need to verify the proposed design that will support development of the simulation's code (Design Verification). A mapping of the proposed design elements back to the conceptual model and your M&S requirements helps to document that your requirements are appropriately addressed and that there is traceability between those requirements and the proposed design.

Before you can verify a conceptual model, you have to have one. In the ideal world, simulation development would not proceed until the underlying M&S requirements were fully identified on the basis of the requirements of the problem at hand and until a fully verified conceptual model was developed from these requirements. In the real world, of course, we all know that M&S development usually proceeds with inadequately defined or rapidly changing requirements. It is very important, however, that you not sacrifice accuracy on the altar of expediency. Take the time to identify your simulation requirements in as much detail as possible early on. Do so by defining your problem concisely and accurately; by defining the simulation outputs, functions, and interactions that will be required to answer your problem; and by specifying, at least in general terms, how much like the real world you need these outputs, functions, and interactions to be. (See Section 1.6.1 for more details.) The developer will then take these requirements and produce a conceptual model.

A *conceptual model* is a simulation developer's way of translating your modeling requirements into a detailed design framework, from which the software that will make up the simulation can be built. A conceptual model typically consists of a description of how your modeling requirements were broken down into model-able pieces, how those pieces fit together and interact, and how they work together to meet the requirements you specified. It also should include a description of the equations and algorithms that will be used to meet your requirements, as well as an explicit description of any assumptions or limitations made or associated with the equations, algorithms, or solution approaches that were used to solve your modeling problem. The conceptual model also should identify how these assumptions and limitations might impact the simulation's ability to meet your requirements, once it is built. The process of reviewing the conceptual model and ensuring that it meets your specified requirements is called *Conceptual Model Verification***.**

After the conceptual model is verified, the developer produces a *software design specification*, which describes exactly how the conceptual model will be translated into software. It defines the components, elements, functions, and specifications that will be used to produce the simulation's software based on the conceptual model. The process of reviewing the detailed design to be sure it conforms to the conceptual model is called *Design Verification*.

Once verified, the conceptual model and its associated design are converted into actual software by the developer. At this point, you have one last verification hurdle to overcome: verification of the software itself (usually called *Code Verification*). Code verification guarantees that the detailed design is implemented correctly in the software. Code verification normally entails detailed desk checking and software testing of the code, comparing it to the design elements, specifications, and operational criteria that were approved during verification of the conceptual model and detailed design, documenting any discrepancies and fixing any problems discovered.

What if you're not building a new simulation, but just want to use an existing one "off the shelf"? How can you determine that the conceptual model and design specifications of this simulation (over which you had no developmental control) meet your M&S requirements? Before we discuss this, let's define what we mean by *off the shelf*.

Most of the models and simulations in this category are called *legacy* M&S because they have some history of prior use. In addition, some legacy models and simulations in wide use were built before the advent and widespread implementation of detailed software design standards and practices. This does not necessarily mean that they are badly designed (although they certainly can be). A good legacy simulation is characterized by a long history of consistent use and development by an active (usually large) user group, good configuration management and documentation, and widely recognized community acceptance of its results.[3] The most important thing that legacy models and simulations may not have that more recent ones do (or, least, should) have, is detailed documentation of their conceptual models and the design specifications that flow from it. Models and simulations without such documentation may require that a suitable substitute for the conceptual model be generated from an analysis of the code as it currently exists and from any available documentation. Once the conceptual model and existing design elements have been identified and documented, however, you still need to determine if the result meets your M&S requirements. Because you had no control over the conceptual model (or the design requirements and specifications) of a legacy model and simulation, the usual verification of the conceptual model and its associated design may not be appropriate. What you can do, however, is review and compare the legacy simulation's assumptions, limitations, and design elements to your M&S requirements to evaluate whether the simulation as it stands meets your requirements. This is called *Conceptual Model Validation* (see below).

It should be clear from the previous discussion that verification requires a clear understanding between you and the simulation developer about your M&S requirements and about the developer's interpretation (and implementation) of those requirements. This understanding and agreement drives the conceptual model, the simulation design and development based on that model, and your ultimate assessment of the simulation's suitability for your application. *Clear requirements and specifications are crucial to cost-effective verification efforts.*

A number of well-established techniques that can be used for verification are discussed in Chapter 4.

1.2.3 Validation Basics

According to DoDD 5000.59, validation is “the process of determining the degree to which a model is an accurate representation of the real world *from the perspective of the intended uses of the model*.” Notice the emphasis. It is critical that the simulation be assessed in terms of how it will be used. Accurate knowledge of how the simulation will be used determines the degree of detail that must be represented for the simulation to provide usable results and the degree of correspondence with real-world phenomena that will be sufficient for you to use the simulation with confidence. The less you really know about how a simulation will be used to solve your problem, the more likely it is that you will have to over-specify validation requirements “just in case.”

Thus, there are two prerequisites for cost-effective validation: a clear understanding of the intended uses of the model, because this sets your requirements for functionality (i.e., what needs to be modeled) and for fidelity (i.e., how well those functions need to match the real world) and a clear definition of the real world. If you don’t have a good definition of what you’re validating against, you won’t be able to determine the difference between a good validation result and a bad validation result. For example, will you validate a simulation against range data, laboratory data, another simulation, or the opinion of experts in the field? Each of these real worlds has inherent drawbacks and limitations that can make or break the apparent validity of a simulation.

In its simplest form, validation consists of comparing a prediction (from a simulation) with an observation (from the real world), and making a judgment about whether the result is good enough for application to your problem. Simple as this concept is, validation techniques are not limited to comparison of simulation results with test data. They also may include sensitivity analyses to test simulation performance against extreme conditions, comparison with other models and simulations known (or assumed) to have validity in the operating range required, and the opinion of subject matter expert (SME) reviews of M&S results.

Validation typically is addressed at two levels: conceptual model validation and results validation. *Conceptual Model Validation* is the determination (usually by a group of SMEs) that the assumptions underlying the proposed conceptual model are correct and that the proposed simulation design elements and structure (i.e., the simulation’s functions, their interactions, and outputs) likely will lead to results realistic enough to meet the requirements of the application. The difference between conceptual model validation and conceptual model verification is a subtle but important one. Conceptual model *verification* ensures that the proposed conceptual model (and its resultant design) satisfies the *functional, interactional, and output* requirements imposed by the specifics of your problem; conceptual model *validation* ensures that the proposed conceptual model (and its resultant design) satisfies the *fidelity, accuracy, or credibility* requirements imposed by the specifics of your problem. The difference is most easily colloquialized as the difference between the questions “Did I build the thing right?” and “Did I build the right thing?”

Results validation compares the responses of the simulation with known or expected behavior from the subject it represents to ascertain that those responses are sufficiently accurate for the range of intended uses of the simulation. This process includes comparison of simulation outputs with the results of controlled tests, sensitivity analyses, or expert opinion.

An important aspect of validation to remember is that validation will not say a simulation is good or bad. It simply measures the difference between simulation outputs and the real world. The user then decides if that difference is small enough for the simulation to be used in a specific application and if the results when used in that application will have the expected accuracy. (More about this in the next section.)

One final observation on validation. Most simulations are composed of thousands of lines of computer code or thousands of electronic circuits and components (or both). The logic diagram of the alternative paths through a typical simulation is extremely large: sufficiently large, in fact, that it is, in practice, impossible to check every possible path. Hence, for all practical purposes, a simulation cannot be completely validated. Therefore, for the question, “Is this simulation validated?” the answer should always be, “Yes, for the conditions specified in the validation report.” Validation is performed on those aspects of a simulation that are important to a particular application. This makes validation feasible and provides the measures of fidelity in areas most important to successful simulation results.

Some of the more common validation techniques and methods are discussed in Chapter 4.

1.2.4 Accreditation Basics

Once a simulation has been verified and validated[4] in accordance with requirements defined by the intended application, an official statement that it is acceptable for the specified use must be made. According to DoDD 5000.59, *accreditation* is “the official certification that a model or simulation is acceptable for use for a specific application.” In many cases, *Expert Review* is the process used to evaluate V&V results in light of M&S requirements defined by the specifics of the problem. These reviews identify credibility gaps, assess their risk to the program, and make recommendations for (or against) accreditation of specific models and simulations.

The accreditation agent (e.g., a program manager) should participate in the earliest stages of M&S development to become familiar with M&S requirements and acceptance criteria and to identify expert review requirements and appropriate SMEs as early as possible. Early involvement helps mitigate the risk of executing an M&S program that will not meet overall program requirements for M&S credibility. In the final stages of the V&V program, the accreditation agent should participate in the summary evaluation of any V&V results and supplemental M&S information to ascertain the adequacy of M&S efforts and the readiness of the M&S suite for final accreditation.

It is important to recognize that accreditation is not (or, at least, should not be considered) a foregone or assumed conclusion. It is a decision that a specific simulation can be used for specific application, based on objective evidence of suitability for the application. Hence, a simulation can receive an accreditation for use in one specific application (e.g., a flight training application) but not be accredited for use in another specific application (e.g., aircraft system design in an acquisition program).

A process leads up to an accreditation decision. This process gathers all the information about specific model or simulation capabilities relative to the requirements of a specific application. This information includes verification and validation results but also includes such things as simulation run time, number of simulation operators required, the simulation’s history of use, documentation status, configuration management, and other factors that will be discussed in Chapter 5.

1.3 Why Do VV&A?

This section offers six reasons why VV&A is a good idea. It’s worth spending a little time dispelling some common misconceptions about the value (or lack thereof) of VV&A. Why all the fuss, anyway? Isn’t VV&A just another check in the box, added to an already lengthy list of such boxes?

In a word, “No.”

This section will discuss six benefits of VV&A:

- Increased confidence in M&S use
- Reduced risk of M&S use
- Increased M&S usability for future applications
- Cost containment
- Potential for better analysis
- Satisfaction of policy requirements

Although by no means an exhaustive list of potential benefits, these have the most impact.

1.3.1 Increased Confidence in M&S Use

A well thought-out program of V&V activities tailored to the application for which a simulation will be used does much to establish or improve confidence in the use of that simulation for that application. V&V increases confidence in models and simulations by providing objective evidence of credibility *within the confines of that intended use*. Notice the emphasis. V&V, by itself, does little to increase confidence in M&S use unless application-specific requirements for credibility are developed and defined for that use. The challenge to the V&V practitioner, therefore, lies in the selection and scoping of that set of V&V tasks most appropriate to the application at hand. Credible tailoring of V&V activities to specific applications, in turn, requires a clear understanding of the contribution that each V&V technique makes to the credibility of M&S and a knowledge of the M&S functions that are critical to the problem at hand.

Chapter 4 defines V&V techniques and their contributions to M&S credibility for specific classes of models and simulations and their applications. Section 1.6.1 discusses tailoring schemes that allow V&V practitioners to focus V&V tasks on the particular requirements of an application to minimize VV&A cost and schedule. As a general rule, however, it is safe to say that the V&V techniques that lend the most credibility to M&S use are not those that cost the most. In particular, V&V status reports and M&S usage histories can help to reduce the scope of new V&V efforts and to indicate the range of applications for which M&S results have been considered acceptable for use. The cost of this aspect of V&V is much less than the detailed code verification and validation with large amounts of test data envisioned by most users when they think of V&V. A history of prior accreditations also lends considerable weight to the choice of models or simulations for a given application by establishing the degree to which M&S results have been considered acceptable by prior users for similar applications. Again, the cost of an accreditation history review is negligible compared to performing more detailed V&V.

1.3.2 Reduced Risk of M&S Use

A major corollary of increased confidence in M&S use is the reduced risk of relying on models and simulations to support major program decisions, objectives, and milestones. Incorrect or inadequate M&S can lead to corrupted system concepts and requirements, poor system design, inaccurate results, negative training, and even system failure, possibly with catastrophic loss.[5] V&V reduces the risk that M&S use will lead to incorrect or indefensible results. The issue in this case is not really "What is the cost of V&V?" but rather "What is the cost of NOT doing V&V?" What is the cost, in terms of time and money, of making an incorrect decision based on M&S results? These hidden costs of avoiding V&V are frequently intangible, unpredictable, and unquantifiable. As a result, they tend to be ignored in the calculation of the value added by V&V. Nevertheless, reduced risk in using M&S is a major benefit of performing V&V *tailored to the application.*

1.3.3 Increased M&S Usability for Future Applications

The requirement to perform V&V to establish the credibility of M&S for use in DoD applications establishes a beneficial dynamic that can reduce the long-term cost of both M&S use and V&V. This is because V&V activities performed by multiple users on a stable simulation, typically one with a well-defined configuration management and development policy (see Section 1.6.2), will, over time, establish a body of evidence supporting its credible use for a wide variety of applications. Different users will, of course, focus their attention on different aspects of V&V to support their individual applications; outside of your program, you have no control over the V&V that gets done. But as the V&V sample space for a specific simulation grows and with it, the body of evidence supporting its credibility, the more likely that it will receive more development and V&V attention. Other models and simulations that perform similar functions but that do not fare well in V&V or that do not have a V&V pedigree adequate to support credible use will give way to those that do. In this way, V&V becomes a natural selection process for the development of fewer models and simulations but with greater capability and established credibility. From this standpoint, your program benefits from the V&V of others for common-use models and simulations. The same dynamic is likely to apply within your own program, meaning that other programs will benefit from your V&V of a particular simulation, just as you benefit from the V&V of others.

Reducing the duplication and improving the credibility of DoD models and simulations may not number among the proximate goals of the typical program manager when V&V is performed. It is clear, however, that the net effect of V&V activity across a spectrum of users of individual models and simulations will be to improve both their capability and their credibility over time.

1.3.4 Cost Containment

If V&V results are documented in a standardized way (see Chapter 6) and if these results are made readily available to the user community, the cost of V&V to support accreditation will drop. New accreditation efforts can build on the V&V results of earlier users. In this way, improvement of the credibility of individual models and simulations becomes a bootstrap process, with multiple users contributing to the body of knowledge about the simulation. This common body of evidence eventually benefits all users of the simulation.

A beneficial consequence of consolidating V&V results across a M&S user community is that V&V becomes market-driven, reducing the duplication of V&V activities. When individual users have to retrace V&V ground that may have been covered by others, the efficiency of overall V&V efforts for the simulation is reduced. But when a consolidated body of V&V knowledge exists, users can focus on the areas of the simulation that need the most attention for their particular application. The analytical needs of a simulation's user community can thus drive the depth to which V&V data are collected, and individual users (like you) in the community no longer waste precious V&V dollars chasing V&V products that already exist.

This assumes, of course, the existence of standard V&V processes and products within individual M&S communities and ready access to this information by individual members of these communities. DMSO is encouraging ready access to V&V information via the MSRR. In this way, both prerequisites for cost-efficient V&V to support accreditation for diverse M&S communities are being met. Your V&V efforts contribute to the body of knowledge about individual models and simulations, and that contribution benefits all users.

1.3.5 Better Analysis

Before widespread use of M&S, effective problem-solving required the clear definition of the problem and its solution objectives, the charting of the analysis with flow diagrams, and the development of an outline of the expected results. With the advent of complex computer simulations that have great predictive power, however, much of the discipline attached to the analytical process has been neglected in favor of understanding the simulation itself. There has been a growing tendency, for example, to focus analytic efforts on gathering valid input data for simulations (see Section 1.6.5), and on taking advantage of the expanded scope of analysis afforded by high-power computers by running a multitude of simulation cases. In essence, analytical depth is being sacrificed for breadth. Rather than being used to do better analysis, models and simulations are being used to do more analysis.

The requirement to perform V&V, however, coupled with the necessity of narrowing its scope to contain costs, can provide an incentive to rejuvenate sound analytical practices within your program. Cost-effective V&V requires the development of detailed M&S requirements that are focused on the intended use of particular models and simulations for particular applications. Development of these requirements necessitates the clear description and full characterization of the analytical problem and approach to identify required information elements, derive appropriate metrics, identify analytical constraints, determine appropriate M&S outputs, and, in general, integrate M&S into your program in a credible way. The discipline required to develop well-defined M&S requirements clarifies analytical issues and facilitates the development of more thoughtful analytic techniques and approaches. Thus, the requirement for cost-effective V&V requires a return to the basic practices of analytical problem-solving that have fallen into disuse. The result can be a tendency to improve the quality of the analysis applied to your program.

This is not to suggest that VV&A automatically leads to better analysis. Improperly done, VV&A can actually detract from simulation credibility by making it appear that critical credibility issues have been addressed adequately, when in fact they have been improperly addressed. It is the synergism and interplay between VV&A and analysis that, when properly managed, can lead to improved confidence in the results of analysis using M&S.

1.3.6 Satisfaction of Policy Requirements

If you’re still not totally convinced of the value of VV&A, there is one more argument that might turn the trick. We’ve held it until last because, although it’s a persuasive argument, it’s not very popular, and it certainly isn’t intellectually satisfying. Simply put, you don’t have much of a choice.

The inescapable facts are these: (a) M&S will be used more and more across DoD (and industry) to save time, money, and resources, and (b) people in *very* high places are *very* worried about how M&S, both new and old, can be integrated into DoD applications in a credible, justifiable, cost-effective way.[6] This means that, like it or not, VV&A will probably play an increasingly influential role in every aspect of DoD operations that contains M&S. And M&S is playing a greater role in every aspect of DoD operations. It's as simple as that.

1.3.7 Benefits Summary

Although the requirement for VV&A of your M&S is going to get harder to address in the coming years, you should have some appreciation by now of why VV&A is worth addressing in the first place. In short, VV&A

- increases the objective confidence you have in your M&S program
- reduces the risk of making the wrong (possibly catastrophic) decision for a critical study, exercise, or acquisition based on incorrect M&S results
- reduces the proliferation of M&S within your program and focuses V&V attention on those models and simulations most useful to your problem
- results can be leveraged to reduce future VV&A costs
- can require the M&S and Analysis shops within your program to focus more on sound analytical practices in order to define the most cost-effective V&V program that meets your requirements for M&S credibility
- meets Service and DoD policy requirements while preserving technical merit

1.4 Where Does VV&A Fit in the Scheme of M&S?

This section offers an overview of VV&A's place in model development and use. It provides a larger context for M&S use in an application.

It should be remembered that M&S is simply a tool or technique that can be used to solve a problem and that VV&A is just a way to gain assurance that the selected model or simulation can produce meaningful results relative to the problem's solution. The problem that needs to be solved is usually called the *application*. The process for solving the problem is usually referred to as the *application process*. The application process context for VV&A and M&S is shown in Figure 1-1.

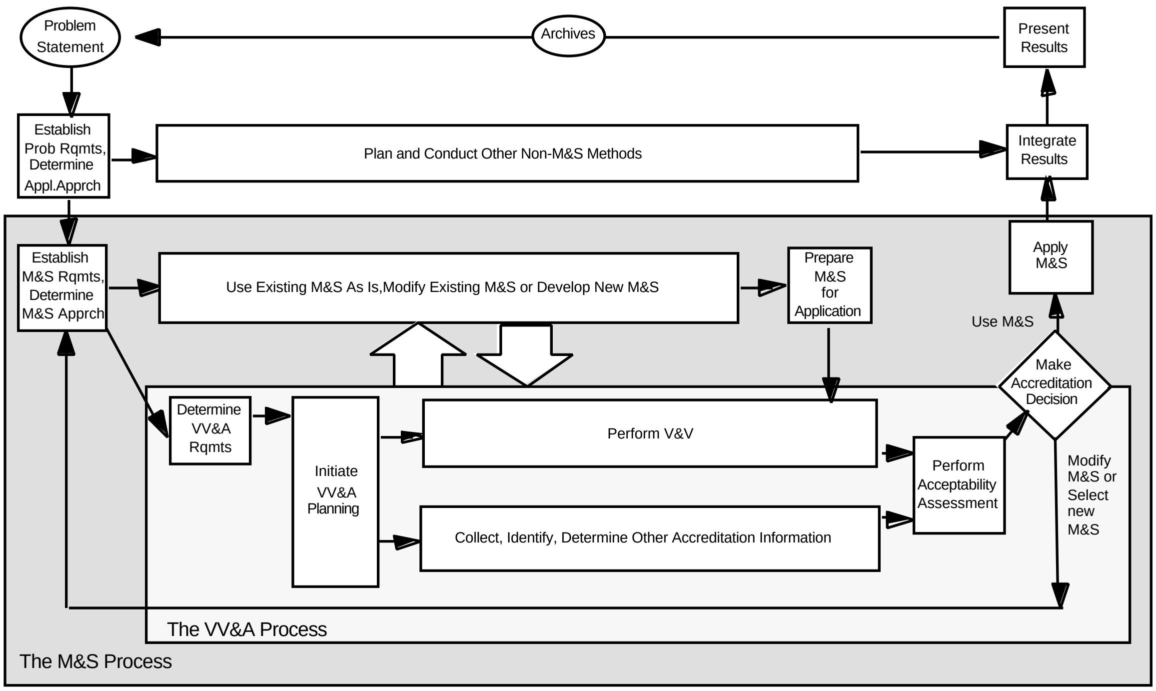

Figure 1-1. VV&A in the Application Process (Scheme of Things)

The application process begins with a clear and unambiguous statement or definition of the problem. A good definition of the problem makes it easier to define its solution requirements. These requirements are the features, characteristics, or functions that are important to the problem and essential to its solution. For example, if the need is to develop a new ECM system, it is essential to define the expected combat systems on which the ECM system will be hosted, the threats against which the ECM system is expected to work, the required effectiveness of the ECM system, the operational environment in which the ECM system will operate, and the other systems expected to be in the operational environment. Measures of Effectiveness (or Merit) that will determine if the requirements have been met are derived from these characteristics. Methods or ways of producing values for these measures or addressing the requirements directly then are determined. These methods can include research into work already accomplished, design analysis, direct testing, or M&S. A complex problem usually employs a number of these types of methods to achieve a robust solution. The set of methods that addresses all the problem requirements is integrated into a consistent, logical application approach.

The application approach shows the problem requirements that will be satisfied by specific methods and the measures that will be used to evaluate the success of each method in fulfilling the solution requirements of the problem. Those methods unrelated to M&S are planned and executed. The requirements to be satisfied by M&S are identified separately and form the basis of the M&S approach. As part of an initial M&S approach, the types of models and simulations that can be used are identified, as well as the criteria for determining when a model or simulation is acceptable for this application. The specific model(s) and simulation(s) to be used for this application are selected according to these criteria. The VV&A status of a model or simulation can be a factor in M&S selection. For complex applications, a number of models and simulations may be necessary to satisfy the M&S requirements of the problem. The M&S approach may call for using specific models or simulations as they are, modifying existing models and simulations, or developing new models or simulations.

Once the M&S suite has been selected and the M&S approach finalized, work can begin on establishing, modifying, or developing the model or simulation. The VV&A process begins immediately and uses the M&S requirements, the acceptability criteria, and the VV&A status of the selected models and simulations to determine the VV&A requirements for this application. Based on these requirements, a plan to accomplish the necessary V&V is developed. Although V&V will produce significant information about the model's or simulation's capability to support the application, additional information beyond V&V is also useful. This other information can include the model's or simulation's configuration management status, documentation status, previous use in other similar applications, and development standards used. This other information (and the V&V results) is a factor in the acceptability assessment. The acceptability assessment compares the model's or simulation's capabilities and limitations to the acceptability criteria and assesses overall its acceptability for this application. This accreditation assessment report includes a recommendation whether to accredit the model or simulation, along with the rationale for that recommendation.

This technical assessment then is given to the accreditation authority, who must decide, using the assessment information provided, whether the M&S suite is acceptable for use in the application. The decision may be to use the M&S suite as it is, to limit the use of the results of the model or simulation, to perform (additional) modifications to the model or simulation, to perform additional V&V, or to reject the M&S suite completely for this application.

If the decision is to use the model or simulation, the M&S runs and exercises are performed. The results are integrated with the non-M&S results to solve the problem.

Archiving the results of the VV&A activity in the appropriate MSRR for future use is important. Any V&V carried out for this application will reduce the amount of V&V that may be necessary for those models and simulations in future applications.

Chapter 5 discusses in more detail the role of VV&A in the context of application problem-solving.

1.5 Common Misperceptions About VV&A

This section describes the limitations of VV&A. It also explains why each new application must be accredited and why V&V must be reviewed (and possibly repeated or expanded) when a model changes. Three common misperceptions about VV&A arise from a misunderstanding of the nature and value of VV&A.

1.5.1 VV&A Is No Substitute for Sound Analysis

VV&A enhances a simulation's credibility and reduces the risk of its use in a particular application, but VV&A cannot guarantee that the M&S results will be correct, that the results will be correctly analyzed and interpreted, or that the right model was chosen to solve the problem. It can identify a model's weaknesses, but the correction of the weaknesses or their workarounds is not a part of the VV&A process. If the M&S requirements or acceptability criteria are incorrect or ill-defined, the likelihood that an incorrect M&S may be selected and used increases. The VV&A process will not assess the correctness of the M&S requirements or acceptability criteria.

The quality of the VV&A process used to support an application also depends on the thoroughness of the VV&A effort and on the capability and experience of the VV&A team. Unfortunately for some applications, VV&A is done in an afternoon meeting of project team analysts who have limited knowledge of specific M&S and application requirements. The results of this kind of VV&A create a higher risk of poor integration of M&S into problem-solving. You get what you pay for.

1.5.2 Accreditation Is Not a One-Size-Fits-All Check in the Box

Accreditation is a decision to use a *specific* simulation for a *specific* application. Each application has a different set of requirements and detailed acceptability criteria. No two problems are exactly alike. V&V can be done without detailed knowledge of the values of simulation acceptability criteria, but accreditation cannot be performed without application-specific requirements and detailed acceptability criteria.

Moreover, when a simulation is modified, it is usually modified to improve its operation, simulation accuracy, or simulation scope. These changes may affect the simulation's suitability for particular applications. The changes to the simulation must be compared with the modeler's intent (verification), and the impact of the changes on simulation output also must be compared with the real-world system or process to measure the increase or decrease in fidelity (validation). Additionally, when the real-world changes or the model or simulation is used for a purpose different from the original intent, previous VV&A results should be reviewed to determine the impact of these changes on the credibility of the simulation. Because the real world is rarely static over any length of time, it is useful to review a model's or simulation's VV&A status periodically to ensure consistency with the current projection of the real world.

The practical impact of all this is that VV&A cannot be considered a solitary task. Although much of the groundwork for accreditation will remain fixed once the basic information is documented during development V&V, accreditation for specific applications (and after simulation changes) is still necessary.

1.5.3 VV&A Is Never Completed

This misperception is really a corollary of the previous one. Many M&S users are surprised when the issue of VV&A activities arises after development or initial accreditation. If you're tempted to say, "I thought we did all that," you have fallen victim to the most common misperception about VV&A.

VV&A is never finished because simulations cannot be verified or validated completely. Complete verification requires testing of every logical branch and condition of the simulation under all possible combinations of input parameters. Complete validation requires comparison of every possible set of input conditions to data run under identical conditions in the real world. It doesn't take a very complex simulation to exceed the number of practically attainable software tests or testable validation conditions.

This does *not* mean, however, that VV&A is an unattainable Holy Grail; it means only that you should expect VV&A activities to continue throughout the life cycle of M&S development *and* application to particular problems. The scope of VV&A required to establish M&S credibility for any particular problem always will be manageable and determined by the specifics of the problem. Ongoing VV&A activities are the price you should expect to pay for ascertaining and maintaining the credibility of your models and simulations.

1.6 Some Practical Considerations

Right now you're probably pretty nervous. VV&A is not just a check in the box, and (oh, by the way) you're never done. The following sections discuss some practical considerations that should put your mind at ease. In particular, we discuss scoping VV&A efforts to meet your requirements; key players, roles, and functions in VV&A to help you organize your efforts; and configuration management issues that help keep track of VV&A activities and relate them to your particular stage of M&S development. All of these topics will help you extend the shelf life of VV&A results.

1.6.1 Scoping and Cost

This section offers guidelines for estimating VV&A needs based on the application type, its importance, and previous VV&A activity. It also discusses how much VV&A is enough.

Right now, you're probably wondering, "What is all this going to cost me? I have heard that software V&V can consume 25–30 percent of my M&S development budget. I don't have 25–30 percent of my budget to devote to *anything*."

Your well-founded concern reflects the recent focus on the credibility of M&S, which has been balanced by an equal concern for the cost of the V&V activities that contribute to it. The M&S community lacks a coherent process that links V&V information to application-specific requirements for M&S credibility. This lack has prevented M&S users (like you) from identifying cost-efficient sets of V&V activities that meet credibility requirements for individual applications. The natural result has been a tendency to overestimate V&V requirements, with the corresponding (mis)perception that "V&V costs too much and takes too long." Operating under this misperception, cost and schedule pressures can lead easily to an irresistible temptation to dilute M&S credibility requirements to meet fixed (usually meager) V&V budgets. The end result leans toward accreditation by fiat, rather than by objective evidence. What's a program manager to do?

1.6.1.1 Exorcising the Cost Demon

First of all, don't be misled by what appear to be overblown estimates of the cost of VV&A. There is a great deal of misinformation on the exaggerated cost of VV&A propagated by people who have little or no first-hand experience in performing it or who have a vested interest in ensuring business continues to be done as usual. The overwhelming evidence from a large number of samples indicates that costs have been well-controlled and tend to cluster or correlate in a predictable manner. Historical data show, for example, that the percentage of M&S development funds devoted to the assessment of M&S credibility spans a reasonably narrow spectrum, from a low of about 5 percent to a high of about 17.5 percent, with most efforts somewhere in the middle range of 10 to 12 percent.

Even these costs depend somewhat on the aspects of VV&A that are included in the estimate. Some think that all V&V and test and evaluation (T&E) activities performed by the developer should be considered part of the total cost of VV&A, leading to the anecdotal estimates of 25–30 percent of development costs, whereas others count only those activities specifically required to accredit a simulation for a given application, leading to estimates closer to half of the previous ones. Either way, the historical record shows that the high estimates tend to include V&V tasks not necessarily essential for M&S accreditation, whereas the minimum levels tend to be a bit Spartan and may not always provide the full range of V&V data necessary to make a strong case for M&S credibility. As in all things, moderation is the key.

1.6.1.2 Trading Off Cost Against Credibility or Risk Reduction

But what constitutes effective moderation? In estimating the costs of your V&V efforts, should you stay closer to 5 percent or 17.5 percent of your M&S budget, or should you just shoot for the average (11.25 percent) and live with the results? How can you tell whether or not the V&V activities you buy for *any* amount of your budget will meet your M&S credibility requirements?

First of all, you'll have to accept that selection of V&V activities on a fixed budget will always involve a trade-off of cost against credibility. Truly cost-effective VV&A seeks to balance the requirement for M&S credibility and risk reduction, driven by the specifics of your application, with real-world constraints, driven by the program M&S budget. Final selection of the exact set of VV&A activities depends strongly on the defined needs, known problem areas, and high-risk aspects of your program, as well as on the availability of tools, methods, human resources, and facilities. When done in good faith, however, VV&A has been shown to provide more in benefits than it costs in resources. It is unquestionably an added-value process, but V&V activities must be chosen correctly. The real question is not, "How much should I spend?" but "What should I buy?"

Guidance from the M&S professional community on how to select the most cost-effective set of V&V tasks to meet a particular requirement for credibility has matured in recent years.[7] The process of selecting V&V tasks rationally within a constrained budget involves answering three key questions about the integration of M&S in your program: What do you need M&S to do? How well do you need M&S to do it? How well do candidate models and simulations do what's needed? If you can answer these three questions, you can select a cost-efficient set of V&V activities that meet your requirements for M&S credibility. Most M&S experts would agree that faithful execution of two activities contributes greatly to the development of a well-focused (hence cost-effective) V&V program.

1.6.1.2.1 Application Analysis. First conduct an in-depth analysis of your problem to define what you want M&S to do. Before *any* decisions about applying M&S to a given problem are made, the problem itself must be defined and articulated clearly enough to see *where* models and simulations help solve the problem and *how* they will help solve the problem. *An ill-defined problem is the most common reason for failure to integrate M&S credibly into program objectives*. A sound problem analysis consists of four elements: (a) a correlation of *clearly articulated* program objectives with the decisions that must be made to reach those objectives, similar to a decision hierarchy or tree; (b) development of a *well-defined* set of Measures of Merit (MOMs)[8] by which each decision will be addressed and resolved; (c) an identification of the program decisions and their associated MOMs that will be addressed, resolved, or supported by M&S; and (d) an identification of the required predictive capabilities that models or simulations must have to support each program decision, i.e., M&S functional requirements. The correlation of program objectives, decisions, MOMs, and M&S functional requirements is the single most important aspect of the V&V tailoring process, because it forms a template for the integration of M&S into your program.

1.6.1.2.2 Acceptance Criteria Definition. Next, develop acceptance criteria for models and simulations you might want to use in your program. Having defined *what* M&S will be required to do (the functional requirements), it remains to determine how *well* candidate models and simulations must do them. The answer lies in two types of acceptance criteria: M&S operational requirements and fidelity requirements.

Operational requirements are nonanalytical requirements, in the sense that they do not contribute to resolution of program decisions or their associated MOMs directly. Instead, these requirements define for example,

- hardware and software requirements, e.g., the models and simulations must run on a certain type of workstation under a certain operating system

- pre- and post-processing requirements for M&S data, e.g., M&S inputs or outputs must be converted to special file formats

- operations and training support requirements, e.g., models and simulations cannot have license agreement or operator training requirements because there is no money or no time for training.

Fidelity requirements are the hardest to define. They state how well required M&S functions (or representations, or entities, as well as the interactions between them) must correspond to the real world (see Section 1.2.3) for the M&S results to be acceptable *for the purpose at hand.*[9] This normally requires the development of a notional "error budget," whereby variations in M&S outputs relate to variations in MOM results, which, in turn, correspond to changes in program decisions. Although it is generally possible to specify the *kind* of V&V that needs to be done to support a given level of credibility (for example, face validation versus results validation), the *amount* of V&V required to establish credibility for a particular application will still depend on a clear understanding of how program decisions are affected by M&S outputs.

1.6.1.3 Selecting V&V Tasks

It is now clear why a precise relationship among program objectives, decisions, MOMs, and M&S is essential. The functional, operational, and fidelity requirements developed by the activities previously described constitute a basic checklist of acceptance criteria with which model and simulation characteristics and capabilities can be compared. This comparison is an essential aspect of V&V tailoring, because it justifies objectively the selection of V&V activities. How is this done?

In a typical legacy M&S case, information on model and simulation capabilities is compiled from available documentation, product literature, existing users, and other sources. It is compared to the functional requirements list to determine if any of the required functions are not modeled. In a typical new M&S case, the functional requirements analysis relies heavily on the planning and requirements documentation and on comparison of the conceptual model with the planned uses of the model or simulation. Information on model or simulation operational characteristics, e.g., how much memory it uses, what programming language it is written in, how long it takes to run a typical case, what hardware and operating system is required, what special training and maintenance is required, is obtained. This information is compared to the operational requirements list to determine if additional resources will be required to maintain and operate candidate models or simulations during their application and to decide whether these additional requirements can or should be met. Finally, the fidelity requirements list is compared to the VV&A histories and current results of the candidate models and simulations to determine the applicability of previous V&V and to identify requirements for additional V&V to address the current problem.

Having identified gaps in the V&V state of functional, operational, and fidelity requirements for candidate models and simulations, a VV&A plan can be developed that prioritizes each gap and describes how it will be addressed using the V&V methods most applicable to each model or simulation. Cost and schedule can be estimated for these tasks based on historical data, and risk assessment and mitigation strategies can be developed depending on the way M&S results affect program decisions.[10]

The payoff for giving faithful attention to these aspects of V&V tailoring is that you now have an audit trail of well-defined program objectives and decisions, M&S acceptance criteria, and V&V data that substantiate the use and acceptance of M&S results.

1.6.1.4 Accounting for Uncertainty

Numerous factors make practical application of these guidelines less than straightforward. For example, V&V program costs can be influenced by the requirement for new or specialized training; long-term site visits at national test ranges to support data collection for model and simulation validation; large capital expenditures for hardware and software; unusual technical efforts requiring significant engineering and analysis; and set-up and maintenance of libraries, data bases, threats files, and the like. The most important aspect of uncertainty, however, is the relationship between the level of V&V required to ascertain the credibility of a model or simulation and the process used to develop its software (called the development *paradigm*).

If you're not going to develop a new simulation for use in your program, i.e., you're going to rely on off-the-shelf or legacy models or simulations, you can more or less skip this section. If you'll be building a simulation for use in your program, however, you'll need to modulate the advice given earlier with the practical realities of model and simulation development described in the following paragraphs.

VV&A must parallel model and simulation development to be truly cost-effective. VV&A planning and execution for models and simulations in development cannot occur without two essential ingredients: (a) optimization of the development paradigm by the M&S developer and (b) a thorough knowledge of the total set of program objectives, requirements, and constraints, which are used to tailor the VV&A approach to the needs of the program. (Yet another reason a clear definition of M&S requirements based on program objectives is essential.) Stated another way, models and simulations can be developed several ways; selection of the best development paradigm is based on the unique set of circumstances, constraints, and application particulars defined by your M&S requirements. This development paradigm can influence VV&A requirements heavily. But, if VV&A is involved early enough in the M&S development cycle, it can have a strong influence on the optimization of the development paradigm.

Several M&S development paradigms are available, from the classic single-pass (or waterfall) approach through the recursive (or evolutionary) approaches that include spiral, prototyping, concurrent engineering, and rapid prototyping variations. The more certainty about the detailed requirements of the new simulation, the more likely that it can be generated successfully and economically using a single-pass development paradigm. This approach assumes that the operational expectations and performance requirements of each model or simulation component (and of the model or simulation overall) and the development environment and infrastructure are reasonably well understood and predictable. As more uncertainty is introduced, the need to iterate (or loop) on problem areas increases. In fact, if you can't define your M&S requirements fully early on, iteration becomes an essential strategy to gain sufficient knowledge to justify proceeding to the next phase of development with reasonable confidence, effectively controlling and managing risk. The complexity of scheduling and budgeting model or simulation development and VV&A activities likewise increases with uncertainty, so that contingency allowances (often called the *management reserve*) must be factored into cost estimates, schedules, and program plans.

Perhaps the simplest way to visualize selection of an appropriate model or simulation development paradigm is to imagine a continuum of certainty and uncertainty. Figure 1-2 lists several of the key attributes that help an M&S developer figure out where the development program belongs along this continuum. The left side of the figure depicts a high degree of certainty about key development factors. This means that the decision makers (you) and the M&S developers have a secure knowledge of and confidence in the technologies, systems and components, similar M&S configurations, communications, protocols, data and data bases, operational requirements, scenarios, and other important data needed to define the model or simulation fully. As long as the requirements are stable and predictable, the waterfall model works well. As the uncertainty about key M&S development factors increases, however, the developer is driven toward iterative development paradigms.

A rapid prototyping paradigm is best used when M&S requirements cannot be defined completely at the beginning of the program. In this approach, part of the model or simulation is built and tested, exercised, or demonstrated to enable the users to work with it and thus help define the next, expanded set of requirements. The process repeats until the user (you) is finally satisfied that the product does all of the essential things. Rapid prototyping is highly adaptive and can be used at will almost any time that a high-risk or unknown part of the model or simulation must be expanded. By building an executable piece of the model that can be demonstrated to the customer and user community (e.g., your M&S shop), feedback and refinement can occur very efficiently. Rapid prototyping is extremely useful in developing and evaluating requirements, proving early design concepts, demonstrating the graphical user interfaces and human interactions, proving critical algorithms, and evaluating the environment and infrastructure. It can be inserted anywhere in the development cycle to help solve technical problems and can be used with virtually any of the other development paradigms.

High Degree of Certainty		High Degree of Uncertainty
• Known technologies		• Unproven technologies
• Known, stable requirements		• Unstable requirements
• Reused, VV&A'ed parts		• Mostly new, untried parts
• Stable design		• Fluid design
• Known communication network		• Undecided communication network
• Predictable performance		• Unknown performance
• Strong tool base		• Sporadic tool application
• Certified data sources		• Indefinite data sources
• Known operational objectives		• Vague operational objectives
• Trained participants		• Nondedicated participants
Waterfall Model	<<<---------------->>>	***Recursive Models***
Minimum V&V	<<<---------------->>>	***Maximum V&V***

Figure 1-2. The Certainty-Uncertainty Continuum

But what's all this got to do with the scope of VV&A activities? Simply put, VV&A activities strongly depend on the development paradigm. Generally, the more uncertainty in M&S requirements, the more effort will be expended on VV&A. It is here that a list of VV&A activities, those that normally would be completed in a comprehensive effort, can be of great help. Because such a VV&A list defines a very rich set of activities, only higher level VV&A efforts will attempt them all. A moderate V&V approach, on the other hand, reduces both the intensity and the number of specific activities planned, focusing on those that are most important to the success of the M&S development program as defined by program requirements. Minimum efforts focus sharply on essential activities.

1.6.1.5 Scoping and Cost Conclusions

It's clear that whether you're using legacy models and simulations or building new ones, defining your M&S requirements based on your specific application is essential to the cost-effectiveness of any VV&A efforts. If you can't (or won't) spend the money to define those requirements, chances are you're going to waste a good portion of whatever V&V dollars you do spend.

VV&A planning should not become a contest to provide the absolute lowest cost effort nor, at the other extreme, to provide more elaborate procedures and analyses than are required. Cost-effective VV&A seeks the best value balance between program needs and real-world constraints. When faced with budgets that appear too low to accomplish the VV&A activities suggested by program requirements for M&S credibility, trade-offs have to be made. These trade-offs should prioritize those activities that have the greatest return on investment (ROI) and that instill and confirm the greatest degree of confidence in the model or simulation. Thus, final selection of VV&A activities must be driven by program particulars: discrete requirements, defined needs, known problem areas, high-risk and critical items, and availability of tools, methods, and key staff.

Tailoring VV&A activities requires careful analysis of M&S requirements, an understanding of the development paradigm (when new models and simulations are being developed), knowledge of problem areas and relevant technologies, knowledge of and access to authoritative data sources, and understanding of the M&S environment and infrastructure. The amount of uncertainty governs the amount of VV&A; that's just common sense. When applied in good faith, as opposed to a desire to check a box, VV&A can add substantial value to the integration of M&S into your program, and its cost can be completely justified by validating the conceptual model, reducing rework, detecting problems early, stabilizing the M&S suite chosen for use, improving analytical efficiency, correlating results, ensuring compatibility, and supporting test and evaluation.

1.6.2 Key Players, Roles, and Functions

This section describes the personnel needed to perform VV&A, the roles and responsibilities of major players, the need for independence in V&V, and trade-offs between independence and ignorance. We explore the appropriate roles of M&S

sponsors, developers, V&V agents, accreditation agents, and accreditation authorities. Before we discuss the substance of this section, however, we have to get a few definitions out the way. According to DoDI 5000.61, the following are the accepted definitions of the terms used in the first sentence of this paragraph:

- *M&S Application Sponsor*—The organization that utilizes the results or products from a specific application of a model or simulation

- *Accreditation Agent*—The organization designated by the application sponsor to conduct an accreditation assessment of an M&S application

- *M&S Developer*—The organization responsible for managing or overseeing models and simulations developed by a DoD Component, contractor, or Federally Funded Research and Development Center[11]

- *Validation (or Verification) Agent*—The organization designated by the M&S application sponsor to perform validation (or verification) of a model, simulation, or federation of models and/or simulations.

In a typical scenario, the application sponsor (the one who needs M&S to solve a problem or answer a question) will designate an accreditation agent, who is responsible for organizing, coordinating, and executing a comprehensive VV&A program that will guarantee the credibility of model and simulation results when used for the sponsor's application. The accreditation agent may further designate a V&V agent who will be responsible for producing the V&V data used to accredit the model, or the agent may act as his or her own V&V agent. The M&S Developer is typically designated by the application sponsor to oversee M&S development activities and to ensure coordination with the V&V agent, but the application sponsor also may retain the duties of M&S Developer. In any case, the exact relationship between these organizational entities can have a bearing on the credibility of the outcome of VV&A activities.

A common (mis)perception holds that V&V must be conducted completely independent of the M&S developer, lest the results be tainted by the demands of advocacy, whence the *I* in IV&V. The M&S developer, however, is (and should be) an essential and integral part of V&V, contributing greatly to its efficiency because the developer is intimately familiar with the design and code details and has been involved in the intricacies of development from the start. The developer understands (or should understand) the requirements best and in the best of cases has maintained close contact with the application sponsor. It is also true, however, that the developer has a vested interest in making the product look good. The need for some kind of independent assessment of the developer’s product seems like a common sense risk reduction strategy.

But there is a down side to independence. Totally independent V&V efforts by the V&V agent can retrace much of the work already done by the M&S developer. Rework is fine if the developer is trusted to provide much of the essential information for the V&V agent. Sometimes, however, the relationship between the V&V agent and the developer can become adversarial, with the V&V agent taking on functions of a Government Inspector General. This opposition can burden the development process with unnecessary baggage that you will ultimately have to pay to carry, all in the name of independence. The question that must be answered is, “How much independence can I afford?”

No real hard and fast rules dictate how much *I* to put in V&V. Some notional V&V roles and responsibilities that worked in the past are shown in Table 1-1, but the final decision must be derived from the trade-off between the M&S budget and the level of confidence and trust that can be placed in the M&S developer. Don’t forget that V&V also may be performed in-house by the application sponsor. Frequently, the M&S developer performs the verification of the model or simulation with V&V agent oversight and assists the V&V agent or the application sponsor during validation.

Table 1-1. Typical VV&A Responsibilities

Activity	***Party***			
	V&V Agent	M&S Developer	Application Sponsor	Accreditation Agent
V&V Acceptability Criteria Report	Assists		Responsible	Assists
Accreditation Plan			Responsible	Performs
V&V Plan	Responsible, Performs	Assists	Uses	Uses
Verification	Responsible	Assists		
Validation	Responsible	Assists		
V&V Report	Responsible, Performs	Assists	Uses	Uses
Acceptability Assessment Report	Assists			Responsible
Accreditation	Assists		Responsible, Performs	Assists
Accreditation Report	Assists		Responsible	Performs

Table 1-1 uses the terms: *Responsible*, *Performs*, *Assists*, and *Uses*. *Responsible* means that the listed party ensures that the specified activity is accomplished. *Performs* means that the listed party carries out the technical work associated with the listed activity. *Assists* means that the listed party helps the responsible or performing party with the activity. *Uses* means the listed party employs the product of the listed activity in performance of some function listed later in the table. Remember that Table 1-1 is only a suggested list of interactions and responsibilities. Ultimately, you must decide how much independence is necessary and affordable.

1.6.3 The Importance of Configuration Management

This section describes the relationship between sound configuration management and cost-effectiveness of VV&A. It provides guidelines for evaluating or implementing configuration management procedures.

Software Configuration Management (C/M) is a development life-cycle process through which the integrity and continuity of software development, upgrades, and maintenance are recorded, communicated, and controlled. C/M can have a profound impact on the sustainability of M&S credibility you have worked so hard to attain through V&V. The key to maintaining the shelf life of V&V work is a structured, workable, and well-maintained C/M process that is integrated with model and simulation development. Because the magnitude of the C/M problem will vary depending on the use of legacy models or simulations, only the most general comments will be given here.[12] It is not an overgeneralization to state, however, that V&V not integrated with C/M will result in repetitive efforts and wasted resources.

But what are the elements of a good C/M process? In general, four major characteristics are the hallmarks of sound C/M practice:[13] (a) a well-defined baseline; (b) standard baseline test cases and data sets; (c) well-defined, coordinated, and supported testing program; and (d) current, thorough documentation. Whether you are using legacy models and simulations or developing your own, you should evaluate the C/M process for these characteristics. If it has all four, you can be reasonably sure that the model or simulation is well-managed and controlled. Some special considerations apply to legacy and new models and simulations, and these are discussed in Chapter 3.

1.6.4 Credibility of M&S Data

All M&S are driven by data, either as direct inputs by the user or as embedded constants that drive simulation characteristics. As perfect as the equations, algorithms, and software design of a M&S may be after conceptual model verification and validation and design verification, it will probably fail results validation if the data that drive the simulation are inaccurate or inappropriate for the task at hand. Data credibility is a major driver of M&S credibility.

But how can the credibility of M&S data be quantified? And what standards should be proposed so that the credibility of data used by multiple users is uniform? Data standards benefit all M&S users by providing increased data credibility, reduced need for data translation, interoperability with the operational community, and M&S reuse. Without data standards, interoperability between models and simulations is much more difficult to achieve. Data definitions common to different systems are needed, definitions that are formal and consistent and that use data standardization policies, procedures, and methodologies.

The M&S community has been wrestling with this issue for some time now, and a final pronouncement of standards, procedures, and guidelines for the certification of data credibility has not been made. Several key concepts have emerged, however. Data verification and validation definitions, processes, and procedures that parallel the M&S definitions, processes, and procedures have emerged. For data, the decision is called *certification* as opposed to *accreditation*. Hence the term data VV&C instead of VV&A. Additionally, data VV&C is viewed from two different perspectives: that of the data producer and that of the data user. The key definitions are as follows:

- Data verification establishes that the data produced conform to the specification.
- Data validation establishes that the data accurately represent the real world.
- Certification establishes that the data are suitable for a specific use.

Currently, each Service tracks data about their models and simulations and stores them at the service's own required level of detail in service-specific format. DMSO has been coordinating an effort to standardize the level of detail, format, and accessibility for all DoD M&S data, including VV&A and data VV&C information. These data will be centrally controlled but accessed in a distributed environment. They will provide information critical to M&S planners and the basis for model and simulation life-cycle management. Additional discussion of data VV&C is in Chapter 3.

1.7 Roadmap to This Guide

Now that you've had an overview of VV&A, you're ready for more detail. Chapter 2 provides a set of governing principles of VV&A based on the experience of Government, industry, and academic experts. Chapter 3 introduces the basics of VV&A processes and sets the context for Chapter 4, which deals with the details of VV&A techniques. Chapter 5 discusses combining V&V information into a sound accreditation decision, and Chapter 6 completes this guide by discussing common reporting formats that will simplify the maintenance of an audit trail of V&V and accreditation support activities.

Endnotes

[1] *V&V* is the term applied to the technical work that supports a decision ("accreditation") to use a model or simulation. The terms are defined later in this section.

[2] For example, a flight training simulator does not need to kill the pilot if he or she crashes the plane during training to accomplish the goal of teaching the pilot not to crash.

[3] Inclusion of the model or simulation in one of the Information Analysis Center (IAC) model repositories can be a good indication of community acceptance of M&S results. IACs are run by the Defense Logistics Agency (DLA) to support a wide variety of DoD analysis needs. DMSO's Modeling and Simulation Resource Repository (MSRR) is, likewise, a good source for M&S resources that are considered authoritative.

[4] *V&V'ed*, in the vernacular, although some object to this casual use of technical terms.

[5] For example, loss of at least one fly-by-wire aircraft has been attributed to M&S inadequacies. See *The Day the Phones Stopped* by L. Lee (Donald I. Fine, Inc., 1991).

[6] See Army Regulation (AR) 5-11 and the associated DA PAM 5-11; Air Force Instruction 16-1001; draft SECNAVINST 5200.1; DoDD 5000.59; and DoDI 5000.61.

[7] VV&A efforts conducted under Military Operations Research Society (MORS), Distributed Interactive Simulation (DIS), and Susceptibility Model Assessment with Range Test (SMART) auspices are notable in this regard.

[8] *MOM* is a generic term that encompasses all measures of value, including Measures of Performance (MOPs), Effectiveness (MOEs), and Outcome (MOOs).

[9] This does not imply an absolute standard of fidelity for all applications but rather a level of fidelity considered good enough. The good must not become the enemy of the best.

[10] Another reason to spend some time defining your program objectives, decisions, and M&S requirements.

[11] For the purposes of this discussion, we will include the organization (government or contractor) responsible for actually building the software under the term *developer*.

[12] See, however, *Configuration Management Requirements Study*, available from the JTCG/AS (JTCG/AS-95-M-005), which discusses DoD and MIL-STDs for software C/M.

[13] Comments on C/M for legacy models and simulations are taken from the study cited in Footnote 12. The goal of the study was to identify common requirements for the C/M of legacy models and simulations, to compare these requirements to current practice, and to make recommendations for improvement.